

The expansion of real forms on the simplex and applications

Yong Yao^a, Jia Xu^{b,*}, Jingzhong Zhang^a

^a*Chengdu Institute of Computer Applications, Chinese Academy of Sciences, Chengdu, Sichuan 610041, China*

^b*College of Computer Science and Technology, Southwest University for Nationalities, Chengdu, Sichuan 610041, China*

Abstract

If n points B_1, \dots, B_n in the standard simplex Δ_n are affinely independent, then they can span an $(n-1)$ -simplex denoted by $\Lambda = \text{Con}(B_1, \dots, B_n)$. Here Λ corresponds to an $n \times n$ matrix $[\Lambda]$ whose columns are B_1, \dots, B_n . In this paper, we firstly proved that if Λ of diameter sufficiently small contains a point P , and $f(P) > 0$ (< 0) for a form $f \in \mathbb{R}[X]$, then the coefficients of $f([\Lambda]X)$ are all positive (negative). Next, as an application of this result, a necessary and sufficient condition for determining the real zeros on Δ_n of a system of homogeneous algebraic equations with integral coefficients is established.

Keywords: simplex, system of algebraic equations, real zeros

2000 MSC: 13J25, 13J30, 16Y60, 68T15, 26D05

1. Introduction

The standard simplex is denoted by $\Delta_n = \{(x_1, \dots, x_n)^T \mid x_i \geq 0, \sum_i x_i = 1\}$. If n affinely independent points $B_1, \dots, B_n \in \Delta_n$, then they span an

[☆]This work was supported by the National Basic Research Program of China (973 Program: 2011CB302402), China National Natural Science Foundation (11001228), and the Fundamental Research Funds for the Central Universities, Southwest University for Nationalities (12NZYTH04).

*Corresponding author.

Email addresses: yaoyong@casit.ac.cn (Yong Yao), j.jia.xu@gmail.com (Jia Xu), zjz101@yahoo.com.cn (Jingzhong Zhang)

$(n-1)$ -simplex denoted by $\Lambda = \text{Con}(B_1, \dots, B_n)$, that is,

$$\Lambda = \text{Con}(B_1, \dots, B_n) = \{\lambda_1 B_1 + \dots + \lambda_n B_n \mid \sum_{i=1}^n \lambda_i = 1, \lambda_1, \dots, \lambda_n \geq 0\}.$$

Λ corresponds to an $n \times n$ matrix $[\Lambda]$ whose columns are B_1, \dots, B_n . Conversely, consider the non-negative matrix $M = [b_{ij}]$, where $B_j = (b_{1j}, \dots, b_{nj})^T$ ($j = 1, \dots, n$). If the determinant $|M| \neq 0$, and the sum of the elements in each column is 1 (i.e., $\sum_{i=1}^n b_{ij} = 1$), then M corresponds to an $(n-1)$ -simplex $\text{Con}(B_1, \dots, B_n)$ in Δ_n , which is briefly denoted by $\text{Con}(M)$.

Given an $(n-1)$ -simplex $\Lambda \subseteq \Delta_n$ and a real form (homogeneous polynomial) $f \in \mathbb{R}[x_1, \dots, x_n]$, we call $f([\Lambda]X)$ the expansion of f on the simplex Λ . This simple expansion has a surprising consequence. Here we give an example.

Example 1 (see [8]) Let $(x, y, z)^T \in \Delta_3$. Prove that

$$f(x, y, z) = 3^7(y^4 z^4 (y+z)^4 (2x+y+z)^8 + x^4 z^4 (x+z)^4 (x+2y+z)^8 + x^4 y^4 (x+y)^4 (x+y+2z)^8) - 2^8(x+y+z)^8 (x+y)^4 (x+z)^4 (y+z)^4 \geq 0.$$

To prove this, first consider the following six matrices:

$$M_1 = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} \end{bmatrix}, \quad M_2 = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{3} \end{bmatrix}, \quad M_3 = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} \\ 1 & \frac{1}{2} & \frac{1}{3} \end{bmatrix},$$

$$M_4 = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{3} \\ 1 & \frac{1}{2} & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} \end{bmatrix}, \quad M_5 = \begin{bmatrix} 0 & 0 & \frac{1}{3} \\ 1 & \frac{1}{2} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{3} \end{bmatrix}, \quad M_6 = \begin{bmatrix} 0 & 0 & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{3} \\ 1 & \frac{1}{2} & \frac{1}{3} \end{bmatrix}.$$

It is easy to see that each M_i corresponds to a 2-simplex Λ_i in the following Figure 1 (e.g., M_1 corresponds to Λ_1 and M_2 corresponds to Λ_2).

Hence we have that

$$\begin{aligned} (\forall X \in \Delta_3) f(X) \geq 0 &\iff (\forall X \in \Lambda_i) f(X) \geq 0, \quad i = 1, \dots, 6, \\ &\iff (\forall X \in \Delta_3) f(M_i X) \geq 0, \quad i = 1, \dots, 6. \end{aligned}$$

Expand the form $f(M_i X)$ with Maple (or Mathematica), we will find that all the coefficients of $f(M_i X)$ are real non-negative numbers. This shows that $(\forall X \in \Delta_3) f(X) \geq 0$.

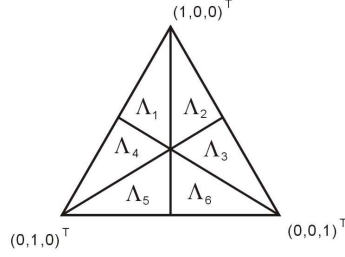


Fig.1. The barycentric subdivision of Δ_3

From the above example, we know that $f([\Lambda_i]X)$ ($i = 1, \dots, 6$), the expansion of f on the simplex Λ_i , can help us to determine the non-negativity of f on the standard simplex Δ_3 . This kind of expansion also has other applications. This work will focus on the expansion of real forms and applications.

Before further illustrating the main results, we will need some notations.

Let $A = (a_1, \dots, a_n)^T$ and $B = (b_1, \dots, b_n)^T$ be two points of Δ_n . Then the ∞ -norm distance $d_\infty(A, B)$ is

$$d_\infty(A, B) = \max_{1 \leq i \leq n} \{|a_i - b_i|\}.$$

Furthermore, for $(n-1)$ -simplex $\Lambda = \text{Con}(B_1, \dots, B_n)$, the diameter $\phi(\Lambda)$ is

$$\phi(\Lambda) = \max_{A, B \in \Lambda} \{d_\infty(A, B)\}.$$

It is well known that (see [1, 3])

$$\phi(\Lambda) = \max_{1 \leq s, t \leq n} \{d_\infty(B_s, B_t)\}, \quad \phi(\Delta_n) = 1.$$

\mathbb{N} and \mathbb{R} are respectively used to refer to the set of all natural numbers and the set of all real numbers. Given $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$, we set $|\alpha| = \alpha_1 + \dots + \alpha_n$ and $\alpha! = \alpha_1! \cdots \alpha_n!$, and for a vector $X = (x_1, \dots, x_n)^T$, we have $X^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. A form (homogeneous polynomial) f of degree d in \mathbb{R}^n is given by

$$f(X) = \sum_{|\beta|=d} \left(\frac{d!}{\beta!} \right) C_\beta X^\beta, \quad \beta \in \mathbb{N}^n, \quad C_\beta \in \mathbb{R}. \quad (1)$$

The minimum value of $f(X)$ over Δ_n , the maximum value of the coefficients of absolute value $|C_\beta|$ and the combination are denoted by λ_f , L_f , and C_n^m

respectively. Namely,

$$\lambda_f = \min_{X \in \Delta_n} \{f(X)\}, \quad L_f = \max_{|\beta|=d} \{|C_\beta|\}, \quad C_n^m = \frac{n!}{m!(n-m)!}. \quad (2)$$

The main results of this paper is the following two theorems.

Theorem 1. Let f be in the form of (1), and let L_f , C_n^m be in the form of (2). The diameter of $(n-1)$ -simplex $\Lambda(\subseteq \Delta_n)$ is denoted by $\phi(\Lambda)$. For a point $P \in \Delta_n$ satisfying $f(P) > 0$ (< 0), if $(n-1)$ -simplex $\Lambda(\subseteq \Delta_n)$ contains the point P and $\phi(\Lambda)$ satisfies

$$\phi(\Lambda) < (dd!C_{n+d-1}^d L_f)^{-1} |f(P)|,$$

then the coefficients of $f([\Lambda]X)$ are all positive (negative).

Theorem 2. Let $Z = \{X \in \Delta_n \mid f_1 = \dots = f_k = 0\}$, where $f_1, \dots, f_k \in \mathbb{Z}[x_1, \dots, x_n]$ are all homogeneous polynomials with degree d_1, \dots, d_k respectively. Denote $d = \max\{d_1, \dots, d_k\}$. Let

$$F = f_1^2 \left(\sum_{i=1}^n x_i \right)^{2(d-d_1)} + \dots + f_k^2 \left(\sum_{i=1}^n x_i \right)^{2(d-d_k)},$$

where H is the maximum of absolute values of coefficients of F , and $\tilde{H} = \max\{H, 4n+2\}$. Then the set Z is not empty if and only if there is an $(n-1)$ -simplex $\Lambda \subset \Delta_n$ such that the coefficients of the polynomial $F([\Lambda]X)$ are not all positive whenever the diameter of Λ satisfies

$$\phi(\Lambda) < (2d(2d)!C_{n+2d-1}^{2d} L_F)^{-1} (2^{4-\frac{n}{2}} \tilde{H} (2d)^n)^{-n2^n (2d)^n}.$$

Here L_F is a notation in (2).

Note: In Theorem 1, "coefficients are all positive" means that every monomial of degree d appears with a strictly positive coefficients. For example, the coefficients of the form $x^2 + xy$ are not all positive, since the coefficient of the monomial y^2 is 0.

2. Proof of Theorem 1

To get started, let us look at the following lemma.

Lemma 2.1. Let f be in the form of (1), and let L_f be in the form of (2). For $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$, the partial derivative of f is denoted by $D^\alpha f$, that is,

$$D^\alpha f = \frac{\partial^{\alpha_1 + \alpha_2 + \dots + \alpha_n} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}.$$

Then we have

$$|D^\alpha f(X)| \leq d! L_f \quad (X \in \Delta_n, |\alpha| \leq d).$$

PROOF. Construct the polynomial

$$F(X) = L_f(x_1 + \dots + x_n)^d - f(X).$$

Since the coefficients of $F(X)$ are all non-negative, then the coefficients of $D^\alpha F(X)$ ($|\alpha| \leq d$), too, are all non-negative, that is,

$$D^\alpha F(X) = D^\alpha(L_f(x_1 + \dots + x_n)^d) - D^\alpha f(X) \geq 0, \quad X \in \Delta_n.$$

Hence

$$|D^\alpha f(X)| \leq d! L_f(x_1 + \dots + x_n)^{d-|\alpha|} = d! L_f \quad (X \in \Delta_n, |\alpha| \leq d). \quad \square$$

Next we will give the proof of Theorem 1.

PROOF OF THEOREM 1. The basic strategy is to use the complete Taylor's formula.

Suppose that $(n-1)$ -simplex $\Lambda = \text{Con}(B_1, \dots, B_n) = \text{Con}([b_{ij}]) \subseteq \Delta_n$, $1 \leq i, j \leq n$. Let $f(P) > 0$ for the point $P = (p_1, \dots, p_n)^T \in \Lambda$, and let $\varepsilon_{ij} = b_{ij} - p_i$, $\varepsilon_{\max} = \max\{|\varepsilon_{ij}|, 1 \leq i, j \leq n\}$. By the definition of $\phi(\Lambda)$, we know that the ∞ -norm distance from the point P to the vertices B_1, \dots, B_n is less than or equal to the diameter $\phi(\Lambda)$, that is,

$$|\varepsilon_{ij}| = |b_{ij} - p_i| \leq d_\infty(B_i, P) \leq \phi(\Lambda).$$

In particular, $\varepsilon_{\max} \leq \phi(\Lambda)$.

From $b_{ij} = p_i + \varepsilon_{ij}$, it follows that

$$[b_{ij}]X = \begin{bmatrix} p_1|X| + (\varepsilon_{11}x_1 + \cdots + \varepsilon_{1n}x_n) \\ p_2|X| + (\varepsilon_{21}x_1 + \cdots + \varepsilon_{2n}x_n) \\ \vdots \\ p_n|X| + (\varepsilon_{n1}x_1 + \cdots + \varepsilon_{nn}x_n) \end{bmatrix} = P|X| + H,$$

where $H = (h_1, \dots, h_n)^T$, $h_i = (\varepsilon_{i1}x_1 + \cdots + \varepsilon_{in}x_n)$ and $|X| = x_1 + \cdots + x_n$.

Consider Taylor's formula

$$f(P|X| + H) = f(P|X|) + \sum_{k=1}^d \sum_{|\alpha|=k} \frac{D^\alpha f(P|X|)}{\alpha!} H^\alpha, \quad (\alpha \in \mathbb{N}^n). \quad (3)$$

Then we know that (3) is a identity formula for the polynomial f of degree d .

Notice that f is homogeneous, hence by (3) we have

$$\begin{aligned} f([A]X) &= f(P|X| + H) \\ &= f(P)|X|^d + \sum_{k=1}^d \sum_{|\alpha|=k} \frac{D^\alpha f(P)|X|^{d-k}}{\alpha!} H^\alpha \\ &= \sum_{k=1}^d \left(\frac{f(P)}{d} |X|^d + \sum_{|\alpha|=k} \frac{D^\alpha f(P)}{\alpha!} |X|^{d-k} H^\alpha \right) \\ &\stackrel{*}{=} \sum_{k=1}^d \sum_{|\alpha|=k} \left(\frac{f(P)}{dC_{n+k-1}^k} |X|^d + \frac{D^\alpha f(P)}{\alpha!} |X|^{d-k} H^\alpha \right) \\ &= \sum_{k=1}^d \sum_{|\alpha|=k} \left(\frac{f(P)}{dC_{n+k-1}^k} - \frac{|D^\alpha f(P)|}{\alpha!} \varepsilon_{max}^k \right) |X|^d \\ &\quad + \sum_{k=1}^d \sum_{|\alpha|=k} \frac{|D^\alpha f(P)|}{\alpha!} (\varepsilon_{max}^k |X|^k + H^\alpha) |X|^{d-k}. \end{aligned}$$

(Note: the identity $\sum_{|\alpha|=k} 1 = C_{n+k-1}^k$ is used at the above sign $\stackrel{*}{=}$.)

It remains to consider the above expression. We will do it in the following two steps.

First, by lemma 2.1, we have

$$|D^\alpha f(P)| \leq d!L_f.$$

Since

$$\varepsilon_{max} \leq \phi(\Lambda), \quad C_{n+k-1}^k \leq C_{n+d-1}^d \quad (k \leq d).$$

Thus, we have

$$\left(\frac{f(P)}{dC_{n+k-1}^k} - \frac{|D^\alpha f(P)|}{\alpha!} \varepsilon_{max}^k \right) \geq \left(\frac{f(P)}{dC_{n+d-1}^d} - d! L_f \phi(\Lambda) \right) > 0.$$

Further, it is obvious that the coefficients of

$$\frac{|D^\alpha f(P)|}{\alpha!} (\varepsilon_{max}^k |X|^k + H^\alpha) |X|^{d-k}$$

are all non-negative real numbers.

As stated above, we see that the coefficients of $f([\Lambda]X)$ are all positive.

□

Remark 1. For the case of $f(P) < 0$, we just need to make a transformation for $f([\Lambda]X)$:

$$\begin{aligned} f([\Lambda]X) &= f(P|X| + H) \\ &= \sum_{k=1}^d \sum_{|\alpha|=k} \left(\frac{f(P)}{dC_{n+k-1}^k} + \frac{|D^\alpha f(P)|}{\alpha!} \varepsilon_{max}^k \right) |X|^d \\ &\quad - \sum_{k=1}^d \sum_{|\alpha|=k} \frac{|D^\alpha f(P)|}{\alpha!} (\varepsilon_{max}^k |X|^k - H^\alpha) |X|^{d-k}. \end{aligned}$$

Then follow the analogous way of the above proof, we will see that the coefficients of $f([\Lambda]X)$ are all negative when $f(P) < 0$.

Then we present a useful corollary of the theorem 1, which is the basis of the successive difference substitution method (see [7, 8, 9, 10]).

Consider the following $n \times n$ matrix G_n (see [9, 10])

$$G_n = \begin{pmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{n} \\ & \frac{1}{2} & \ddots & \vdots \\ & & \ddots & \frac{1}{n} \\ 0 & & & \frac{1}{n} \end{pmatrix}.$$

We say that a matrix is the barycentric matrix if it can got by permuting the rows of the matrix G_n (when $n=3$, see Example 1).

Suppose that S_n is a symmetric group (permutation group) on the set $\{1, 2, \dots, n\}$. Let P_σ be a permuting matrix corresponding to the permutation σ ($\sigma \in S_n$), and let I be an identity permutation. Then the barycentric matrix can be written as

$$G_\sigma = P_\sigma G_n = P_\sigma G_I, \quad \sigma \in S_n.$$

It is easy to see that the number of the barycentric matrices of order n is $n!$.

Corollary 1. Let $f \in \mathbb{R}[x_1, \dots, x_n]$ be a form of degree d and let λ_f be the minimum value of f on Δ_n . Then:

1. $\lambda_f > 0$ if and only if there is a positive integer N such that the coefficients of the form $f(G_{\sigma_1} G_{\sigma_2} \cdots G_{\sigma_N} X)$ ($\forall \sigma_i \in S_n, i = 1, \dots, N$) are all positive.

2. $\lambda_f < 0$ if and only if there are a positive integer N and a simplex $\text{Con}(G_{\sigma_1} G_{\sigma_2} \cdots G_{\sigma_N})$ such that

$$f(G_{\sigma_1} G_{\sigma_2} \cdots G_{\sigma_N} (1, \dots, 1)^T) < 0.$$

PROOF. (Sketch) To prove this corollary, we need the following three results:

(a) For an arbitrary natural number m , we have that

$$\Delta_n = \bigcup_{\sigma_1 \in S_n} \cdots \bigcup_{\sigma_m \in S_n} \text{Con}(G_{\sigma_1} \cdots G_{\sigma_m}).$$

(b) The diameter of the simplex $\text{Con}(G_{\sigma_1} \cdots G_{\sigma_m})$ satisfies

$$\phi(\text{Con}(G_{\sigma_1} \cdots G_{\sigma_m})) \leq \left(\frac{n-1}{n} \right)^m.$$

(c) There is a positive integer N satisfying

$$\left(\frac{n-1}{n} \right)^N < (dd!C_{n+d-1}^d L_f)^{-1} |\lambda_f|. \quad (4)$$

Here (a) and (b) are the immediate consequence of barycentric subdivision of the simplex Δ_n (see [1, 3]). For a given form f , the right hand of the inequality (4) is a constant whenever its left hand can be arbitrarily close to 0. Hence such N that satisfies the (4) is existent. Then by applying Theorem 1, the corollary follows immediately. \square

3. Proof of Theorem 2

The discussion in this section is restricted in the polynomial ring $\mathbb{Z}[x_1, \dots, x_n]$. The following lemma is essential in proving Theorem 2.

Lemma 3.1. (see [2]) Let $T = \{x \in \mathbb{R}^n \mid f_1(x) = \dots = f_l(x) = 0, f_{l+1}(x) \geq 0, \dots, f_m(x) \geq 0\}$ be defined by polynomials $f_1, \dots, f_m \in \mathbb{Z}[x_1, \dots, x_n]$ with degrees bounded by an even integer d and coefficients of absolute value at most H , and let C be a compact connected component of T . Let $g \in \mathbb{Z}[x_1, \dots, x_n]$ be a polynomial of degree $d_0 \leq d$ and coefficients of absolute value bounded by $H_0 \leq H$. Then, the minimum value that g takes over C is a real algebraic number of degree at most

$$2^{n-1}d^n$$

and, if it is not zero, its absolute value is greater or equal to

$$(2^{4-\frac{n}{2}}\tilde{H}d^n)^{-n2^n d^n},$$

where $\tilde{H} = \max\{H, 2n + 2m\}$.

PROOF OF THEOREM 2. We first show that if the coefficients of $F([\Lambda]X)$ are not all positive, then the set Z is not empty. We will prove this by contradiction. It is clear that $Z = \{X \in \Delta_n \mid F = 0\}$ since $Z = \{X \in \Delta_n \mid f_1 = \dots = f_k = 0\}$. Let $T = \Delta_n = \{(x_1, \dots, x_n)^T \mid \sum_i x_i = 1, x_i \geq 0\}$, and let $g = F$. By Lemma 3.1, if there is no root of F on T , then the minimum value λ_F of F on T satisfies

$$\lambda_F \geq (2^{4-\frac{n}{2}}\tilde{H}(2d)^n)^{-n2^n(2d)^n},$$

where $\tilde{H} = \max\{H, 4n + 2\}$.

Thus, for a simplex Λ , the diameter $\phi(\Lambda)$ satisfies

$$\phi(\Lambda) < (2d(2d)!C_{n+2d-1}^{2d}L_F)^{-1}\lambda_F,$$

since it satisfies

$$\phi(\Lambda) < (2d(2d)!C_{n+2d-1}^{2d}L_F)^{-1}(2^{4-\frac{n}{2}}\tilde{H}(2d)^n)^{-n2^n(2d)^n}.$$

By Theorem 1, we know that the coefficients of the form $F([\Lambda]X)$ are all positive, which contradicts the condition that the coefficients of $F([\Lambda]X)$ are not all positive. Hence Z is not empty. The converse is trivial and the theorem is proved. \square

Finally, we should remark that if you are interested in the other applications about the expansion of real forms on the simplex, please see [4, 5, 6, 7, 8, 9, 10].

References

- [1] M. A. Armstrong, Basic Topology, Springer-Verlag, New York, Berlin, Heidelberg, 1983, pp. 125-127.
- [2] G. Jeronimo, D. Perrucci, E. Tsigaridas, On the minimum of a polynomial function on a basic closed semialgebraic set and application, <http://arxiv.org/abs/1112.0544v1>.
- [3] J. R. Munkres, Elements of Algebraic Topology, Perseus Books Publishing, Cambridge Massachusetts, 1984, pp. 86-88.
- [4] J. Xu, Y. Yao, Completion of difference substitution method based on stochastic matrix, Acta Mathematica Sinica, Chinese Series. 54 (2011) 219-226.
- [5] J. Xu, Y. Yao, Rationalizing algorithm and automated proving for a class of inequalities Involving Radicals, Chinese J. Computer. 31 (2008) 24-31.
- [6] J. Xu, Y. Yao, An algorithm for determining copositive matrices, Linear Algebra Appl. 435 (2011) 2784-2792.
- [7] L. Yang , Solving harder problems with lesser mathematics, in: Proceedings of the 10th Asian Technology Conference in Mathematics, ATCM Inc, 2005, pp. 37-46.
- [8] L. Yang, B. Xia, Automated Proving and Discovering on Inequalities, Science Press, Beijing, 2008, pp. 163-175.(in Chinese)
- [9] L. Yang, Y. Yao, Difference substitution matrices and decision on non-negativity of polynomials, J. Sys. Sci. & Math. Scis., 29(2009) 1169-1177 (in Chinese).
- [10] Y. Yao, Successive Difference Substitution Based on Column Stochastic Matrix and Mechanical Decision for Positive Semi-definite Forms, Scientia Sinica Mathematica. 40 (2010) 251-264 (in Chinese). < <http://arxiv.org/abs/0904.4030v3> >